EXHIBIT A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	:	
Pieczenik	:	Group Art Unit: 1805
Serial No. 07/662,764	:	Examiner: Lebovitz
Filed: February 28, 1991	:	
For: METHOD AND MEANS FOR SORTING AND IDENTIFYING BIOLOGICAL INFORMATION	:	Attorney docket no. 4-89A

DECLARATION OF LORANCE L. GREENLEE UNDER 37 C.F.R.1.131

I, LORANCE L. GREENLEE, hereby declare:

1. That I am a patent attorney currently with the firm of Greenlee and Winner, of 5370 Manhattan Circle, Suite 201, Boulder, Colorado 80303, which firm represents George Pieczenik, the inventor of the above-identified patent application;

2. That between June of 1987 and January of 1992, the firm was called Greenlee and Associates, and that Greenlee and Associates represented George Pieczenik in the prosecution of the parent application of the above-identified patent application;

3. That the cover letter from Roy A. Durham to me, a photocopy of which is submitted herewith marked Exhibit C and the accompanying draft manuscript, a photocopy of which is submitted herewith as

Exhibit B, were received by facsimile in my office at the same time from Roy Alan Durham and the photocopies submitted herewith are true copies of the documents retained in my files, from which dates and irrelevant information have been blacked out;

4. That the letter of Exhibit C bears a date that is prior to August, 1990, and that the letter also bears a "Received" date stamp which is prior to August 1990, and that it was the custom in my office that all letters received were stamped on the date the letters were opened for inspection or on the date when a facsimile transmission was reviewed, and that the letter of Exhibit C was so stamped;

5. That I reviewed the letter of Exhibit C, on the date of receipt or shortly thereafter, and understood that the manuscript accompanying the letter (Exhibit B) contained disclosure of a random oligonucleotide library contained within a bacteriophage expression vector capable of expressing a library of epitopes of random amino acid sequence.

6. That the cover letter written by Roy Durham (Exhibit C) and the manuscript written by George Pieczenik (Exhibit B) have been retained since receipt in the files of Greenlee and Winner; and

7. That the portions of the cover letter marked Exhibit C which is blacked out do not relate to the content of the manuscript marked Exhibit B.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that willful false statements may jeopardize the validity of this application or any patent issuing therefrom.

Date: May 15, 1992

Lorance L. Greenlee

LORANCE L. GREENLEE
Reg. No. 27,894

Exhibit B (re: USSN 07/662,764)

Is the universe of antibody specificities open or closed? If it is closed than this implies that :

A) one can design epitopes for any possible antibody without immunizing for that epitope. A self addressing sort can be used as a screening procedure to identify the proper paired correspondences between antibody and epitope.

B) the universe of epitopes can be specified in some fashion and in principle synthesized.

C) one will find an independently derived antibody that has the same specificity as a previously isolated antibody. This is called the "second hit" experiment and can be used to define the effective specificity population size of antibodies (EPS). This is logically similar to defining a complementation group in genetics.

Though the universe of epitopes may be large, if it is closed, it can be defined by rules, algorithms or an iterative procedure.

Alternatively, if the universe of antibody specificities is open than this implies that:

A) one cannot design an epitope a priori to immunization.

B) the universe of epitopes cannot be specified or synthesized.

C) one should not be able to find an independently derived antibody that has the same specificity as a previously isolated antibody.

One class of epitopes are linear peptides sequences 5 amino acids in length. The known crystal structure of Fab fragment and lysozyme ( ) show that contact points on lysozyme to not span more than 5 amino acids, even though there are two such points widely spaced. Earlier work ( )on antibody binding of carbohydrates and from glycosidase cleavage show that 5-6 sugar lengths are protected from glycosidase cleavage. The frequency of finding antibodies to substrates not used in the immunization process suggest that the antibody binding site is small. Peptide competition experiments, called epitope mapping ( ), also suggest that peptide sequences 4-5 amino acids long can compete for antibody binding with specificity. In addition, linear sequences differing by one amino acid can mimic these bindings with various degrees of specificity ( ).

It should be pointed out that five amino acid lengths is a length of amino acid sequence that can bind with differential specificity to an antibody. This is not necessarily the size or the complete composition of the immunizing entity. The operational relationship between the immunizing entity and the binding entity can only be

resolved when an in vitro immunization-dependent antibody synthesis system is developed.

What is the universe of epitopes that are linear peptides five amino acid long?

The universe of epitopes that are linear peptides five amino acid long cannot be larger that 20 to the power 5, or 3,200,000 different epitopes.

Can one synthesize the universe of all possible linear peptide epitopes five amino acid long?

This can be done two fundamentally different approaches. One approach is to synthesize each of the possible 3.2 million peptide pentamers separately and sequentially. This would be a logical extension of the epitope mapping strategy except that the immunizing sequence is unknown and all pentamers have to be synthesized.

An alternate strategy is to randomly synthesize peptides and separate each from the other. While the chemical synthesis of random peptides is not difficult the fractionation of 3.2 million pentamers from each other is difficult. A way around this fractionation problem, is to synthesis the coding for each peptide sequence in some expression vector system and fractionate each clone from its neighbor.

On Generating the Universe of Pentapeptide Epitopes by Random Nucleic Acid synthesis:

One way to generate each possible pentapeptide sequence is to take advantage of the fact that a random distribution of nucleotide sequences that are fifteen nucleotides long can code for the complete random distribution of peptide sequences five amino acids long. The random distribution of peptide sequences should contain at least one member of the population of 3.2 million peptide sequences if it is large enough ( ).

Because the genetic code is degenerate i.e. there are 61 codons coding for 20 amino acids; each amino acid, on the average, has 61/20 or 3.05 codons coding synonymously ( ). In terms of the nucleotide universe, there are 61 to the power 5 possible nucleotide sequences coding for the 3.2 million pentameric epitopes. Therefore, there are 844,596,301 possible nucleotide sequences coding for 3,200,000 possible pentapeptide sequences. This means that there are 263.94 synonymous codings for each pentapeptide

sequence. This high degree of synonymous degeneracy allows us one way of evaluating whether one has generated the universe of possible pentameric epitopes. Generating 3-5 synonymous representations of the coding for the pentapeptide universe statistically suggests an almost complete representation of each member of the pentameric universe. That is, if the nucleotide distribution generated is equimolar and random, one would expect that if one randomly generated 3-5 synonymous codings for any particular pentameric peptide sequence, one would have had a statistically good chance of having generated any other pentameric peptide sequence in the population of 3.2 million possible pentamers.

Experiments to Demonstrate Generating a Random Distribution of Nucleotide Sequences Coding for Pentameric Peptide Sequence

In order to create a random distribution, and, thereby, a least one copy of each of the possible pentapeptide sequence I decided upon consultation with Fred Sanger to synthesize a random distribution of fifteenmers. The strategy to synthesize rather than degrade was arrived at because Sanger was in his nucleotide synthesis phase of his career rather than degradative phase. Needless to say, one could in principle degrade DNA down to 15 nucleotides and add linkers and then insert this random population into an expression vector.

The lab had not done extensive synthesis with mixed nucleotide substrates and though reported in the literature it was not clear whether one could synthesize a random distribution of mixed oligomers and test the product. In addition, given the chemical synthesis of one strand how does one generate the complementary strand to create the necessary double stranded DNA piece necessary for inserting into the expression vector.

A nucleotide oligomer containing the following sequence was synthesized:

GATCTTNNNNNNNNNNNNNNNAA where N is equal molar amounts of G,A,T, and C. That is 4 to the power 16 or 4,294,967,296 different DNA molecules were synthesized. Because I added an AA sequence at the 3 prime end and a TT sequence at the 5 prime end, this sequence can base pair with itself in phase on both strands if GAT is the sense phase. In addition, the sequence was synthesized to be able to be ligated in a Bam site (GGATCC) with out regenerating a Bam site in order to use Bam nuclease selection against the parental vector.

One test of the randomness of the synthesis is that half of the 4.3 billion sequences should be able to base pair with the other half. The sequences was purified on a Sep pak column, lyphilized, and resuspended in ligation buffer then brought up to 100 ^{o} C for 5 minutes and brought down to room temperature slowly overnight. This was then ligated into a Bam cleaved f1 bacteriophage RF purified from agarose. This ligation product was then transfected into freshly prepared competent TG1 cells and plated. Representative plaques were picked and screened using only one sequencing track to identify those with inserts. About one/third of Bam re-cleaved ligated RF generated phage containing inserts. These were then plaque purified and sequenced. Figure 1 shows the first of these randomly generated sequences. This was important because it confirmed that 1)it was possible to synthesize a random equimolar sequence 16 nucleotides long, 2) that a sequence could find its complement or one sufficiently similar and 3) align itself in the proper phase in order to be properly ligatable and readable when expressed as a fusion peptide.4) In addition, determining a sequence from the phage DNA showed the insert was stable and purifiable and the phage was viable (no matter what some of my former students thought would happern)

Table 1 lists twenty phage containing DNA inserts coding for pentamers that were generated using this random synthesis method. The table shows that the total composition of the inserts reflect the original equimolar synthesis composition. The peptide sequences coded by these random DNA sequences are also found in the data base or protein sequences at a frequency expected for a codon distribution determined by random nucleotide sequences ( ).

We have created the universe of pentapeptide sequences using randomly synthesized DNA and expressed these coding sequences on the surface protein of a bacteriophage. This allowed the purification and sequencing of each individual pentapeptide sequence.

1)
TCTTTACCAGTCCGCTCGGTAAGATCCTCA
TGAGGATCTTACCGAGCGGACTGGTAAAGA
THRGLUARGTHRGLY LYS
T E R T G K
Phaseolin-Kidney bean

2)
TCTTGTATCGAGTCTTGCATAAGATCCTCA
TGAGGATCTTATGCAAGACTCGATACAAGA
METGLN ASPSER ILU GLN
M Q D S I Q

3)
TCTTCGCCCTCTGACCCCGCAAGATCCTCA
TGAGGATCTTGCGGGGTCAGAGGGCGAAGA
ALAGLYSER GLUGLY GLU
A G S E G E

4)
TCTTGCTTCGGAAATATCTGAAGATCCTCA
TGAGGATCTTCAGATATTTCCGAAGCAAGA
GLN ILU PHEPROLYS GLN
Q I F P K Q

5)
TCTTGCCGTTGGAGGATGTTAAGATCCTCA
TGAGGATCTTAACATCCTCCAACGGCAAGA
ASN ILU LEU GLNARGGLN
N I L Q R Q

Fibrinogen gamma B chain precursor
Fibrinogen gamma A chain precursor

6)
TCTTTGAGTTTCAGCGATGGAAGATCCTCA
TGAGGATCTTCCATCGCTGAAACTCAAAGA
PROSER LEULYS LEU LYS
P S L K L K
P3 protein- Bluetongue virus
H-2 class 1-related secreted histocompatability

7)
TCTTGGAGCGCCCTCGGTGTAAGATCCTC
TGAGGATCTTACACCGAGGGCGCTCCAAGA
THRPROARGALALEU GLN
T P R A L Q
RNA -directed RNA polymerase

8)
TCTTTGCCCACGAATTCTAGAAGATCCTCA
TGAGGATCTTCTAGAATTCGTGGGCAAAGA
LEUGLU PHE VALGLYLYS
L E F V G K

9)
TCTTGCCTGTCGAGCACGCTAAGATCCTCA
TGAGGATCTTAGCGTGCTCGACAGGCAAGA
SERVAL LEUASPARGGLN
S V L D R Q
Coat protein-Cauliflower mosaic virus
Anthranilate synthase

10)
TCTTGATGTACTTTGTCTTGAAGATCCTCA
TGAGGATCTTCAAGACAAAGTACATCAAGA
GLN ASPLYS VALHIS GLN
Q D K V H Q
Beta casein-bovine

11)
TCTTCTGCTTGATATACTTCAAGATCCTCA
TGAGGATCTTGAAGTATATCAAGCAGAAGA
GLUVAL TYRGLN ALAGLU
E V T Q A E

Nucleocapsid protein N-Punta Toro phlebovir
Tyrosine amino transferase-rat

12)
TCTTCGGGAGTAAGGAAAACAAGATCCTCA
TGAGGATCTTGTTTTCCTTACTCCCGAAGA
VALPHELEUTHR PRO GLU
V F L T P E
Pol polyprotein-Bovine leukemia virus

13)
TCTTTGTTGGTTATGTATAGAAGATCCTCA
TGAGGATCTTCTATACATAACCAACAAAGA
LEUTYRILU THR ASN LYS
L Y I T N K

14)
TCTTTTCCTATATCCGCGTCAAGATCCTCA
TGAGGATCTTGACGCGGATATAGGAAAAGA
ASPALAASPILUGLY LYS
D A D I G K

1,680,178 RESIDUES IN DATA BASE. 20**5 OR 3,200,000 POSSIBLE PENTAPEPTIDE SEQUENCES. THEREFORE EXPECT 1,680,178/3,200,00=.5 TIMES 14 OR 7 HITS. WE HAVE 8 HITS, WITH 6 HITTING 0 TIMES, WITH 4 HITTING 1 TIME AND 4 HITTING 2 TIMES.

Both the nucleotide composition and the amino acid search frequencies suggest a random distribution has been achieved and therefore the whole population is represented in a sufficiently large sample of phage i.e. 10**9 phage or about .1ml of phage.

Base Composition Analysis of Randomly Synthesized Coding for Epitopes

1) C T T A C C G A G C G G A C T G G T A A A
2) C T T A T G C A A G A C T C G A T A C A A
3) C T T G C G G G G T C A G A G G G C G A A
4) C T T C A G A T A T T T C C G A A G C A A
5) C T T A A C A T C C T C C A A C G G C A A
6) C T T C C A T C G C T G A A A C T C A A A
7) C T T A C A C C G A G G G C G C T C C A A
8) C T T C T A G A A T T C G T G G G C A A A
9) C T T A G C G T G C T C G A C A G G C A A
10) C T T C A A G A C A A A G T A C A T C A A
11) C T T G A A G T A T A T C A A G C A G A A
12) C T T G T T T T C C T T A C T C C C G A A
13) C T T C T A T A C A T A A C C A A C A A A
14) C T T G A C G C G G A T A T A G G A A A A

T) 01414	0 4 1 3 5 0 4 7 4 1 3 2 0 3 2	0	= 39
C) 14 0 0	5 4 4 2 3 4 5 1 4 3 6 2 5 2 6	6	= 56
G) 0 0 0	4 1 3 7 1 6 2 2 3 5 0 5 5 6 3	3	= 53
A) 0 0 0	5 5 6 2 5 4 3 4 3 5 5 5 4 3 3	5	= 62
Totals	14 x 15 = 210		210

Coding Strand Composition — Phage Strand(+) Composition

T = 39/210 = 18.6% = A
C = 56/210 = 26.7% = G
G = 53/210 = 25.2% = C
A = 62/210 = 29.5% = T

T	25 - 18.6 = 6.4 X 6.4 = 40.96/25	=	1.64
C	25 - 26.7 = -1.7X-1.7 = 2.89/25	=	.12
G	25 - 25.2 = -.2 X -.2 = .04/25	=	.04
A	25 - 29.5 = -4.5X-4.5=20.25/25	=	.81

X^2 = 2.57 ; 3 D.F.

A distribution with a Chi-square of 2.57 and 3 degrees of freedom can be gotten randomly 50% of the time. Therefore, our observed distribution does not differ significantly from our expected (and synthesized) global base composition.

BINGO BIOTECH

MEMO [redacted]

TO: LORANCE L. GREENLEE, ESQ.

DEAR LORANCE:

PLEASE FIND ACCOMPANYING DRAFT #2 FAIRLY HOT OFF THE PRESSES. AS YOU'LL SEE, IT'S NO SMALLER IN SCOPE--BUT SEEMINGLY THIS IS THE MANDATE FROM SIR AARON I.E., GET CREDIT FOR THE CONCEPT, EVEN BETTER IF IT'S LARGE.

IF I CAN FIND A MAC IN NEW YORK, WE CAN WRAP THIS UP FAIRLY QUICKLY NOW.

BEST,

ROY ALAN DURHAM
CHAIRMAN & CEO

RAD/JP
ENCL.

Exhibit C (re: USSN 07/662,764)

243 South Muirfield Road Los Angeles, CA 90004 *Phone:* (213) 939-3935 *Fax:* (213) 939-2144